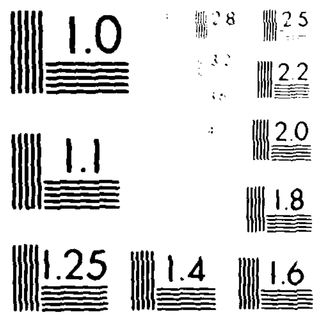
AD-A087 696 CALIFORNIA UNIV BERKELEY STATISTICAL LAB F/G 12/1
LINEAR ESTIMATORS OF MEAN VECTOR IN LINEAR MODELS: PROBLEM OF A--ETC(U)
JUL 80 W KLONECKI N00014-75-C-0159
UNCLASSIFIED CU-SL-80-05-ONR NL

1 OF 1
AD A087696

END
DATE
FILMED
9-80
DTIC



1.0
2.8
2.5
3.2
2.2
2.0
1.1
1.8
1.25
1.4
1.6

LINEAR ESTIMATORS OF MEAN VECTOR IN LINEAR MODELS: PROBLEM OF ADMISSIBILITY

Witold Klonecki

July 1980

STATISTICAL LABORATORY

University of California

BERKELEY

12

LINEAR ESTIMATORS OF MEAN VECTOR IN LINEAR MODELS: PROBLEM OF ADMISSIBILITY

W. Klonecki

Statistical Laboratory
University of California, Berkeley, California 94720

and

Mathematical Institute
Polish Academy of Sciences, Warsaw, Poland

ABSTRACT

Let Y be an n-variate random vector with expectation $\theta \in R^n$ and covariance matrix $V \in \mathcal{V}$, where $\mathcal{V}$ is assumed to be a closed convex cone of non-negative definite matrices of order $n \times n$. It is desired to characterize admissible estimators of θ among linear estimators LY, where L is an $n \times n$ matrix. The main result states that in the case when $\mathcal{V}$ coincides with the set of all non-negative definite matrices LY is admissible if and only if the eigenvalues of L are in the closed interval $[0,1]$. Necessary and sufficient conditions for admissibility when $\mathcal{V}$ consists of positive definite matrices are also given. This latter result generalizes the well-known theorem of Cohen on characterization of admissible linear estimators within models where $\mathcal{V}$ is generated by the unit matrix.

14 CU SL-70-05-ONR

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
ONR 80-05

2. GOVT ACCESSION NO.
AD-A087696

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
(6) Linear Estimators of Mean Vector in Linear Models: Problem of Admissibility.

5. TYPE OF REPORT & PERIOD COVERED
(9) Scientific rept.

6. PERFORMING ORG. REPORT NUMBER
ONR 80-05

7. AUTHOR(s)
(10) Prof. Witold Klonecki
Polish Academy of Sciences
Wroclaw, Poland

8. CONTRACT OR GRANT NUMBER(s)
(15) ONR N000 14-75-C-0159

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Statistical Laboratory
Department of Statistics
University of California, Berkeley, CA 94720

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
Washington, D.C. 20014

(11)

12. REPORT DATE
Jul[redacted]80

13. NUMBER OF PAGES
56

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
(12) 59

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
This document has been approved for public release; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
convex cone
covariance matrix
eigenvalue
risk function

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let Y be an n-variate random vector with expectation $\theta \in R^n$ and covariance matrix $V \in \mathcal{V}$, where $\mathcal{V}$ is assumed to be a closed convex cone of non-negative definite matrices of order $n \times n$. It is desired to characterize admissible estimators of θ among linear estimators LY, where L is an $n \times n$ matrix. The main result states that in the case when $\mathcal{V}$ coincides with the set of all non-negative definite matrices LY
(CONT.)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (CONT.)

is admissible if and only if the eigenvalues of L are in the closed interval [0,1]. Necessary and sufficient conditions for admissibility when V consists of positive definite matrices are also given. This latter result generalizes the well-known theorem of Cohen on characterization of admissible linear estimators within models where V is generated by the unit matrix.

I. THE MODEL

Let Y be an n-variate random vector with expectation $\theta = EY$ and covariance matrix $V = \text{cov}(Y)$. The parameters are (θ,V). Denote the parameter space by $\mathcal{P}$, which is assumed to be of the form $R^n \times \mathcal{V}$, where $\mathcal{V}$ is a closed convex cone of non-negative definite matrices.

Call LY, where L is an $n \times n$ matrix, a linear estimator of $\theta = EY$. The risk function used is

$$R(L|\theta,V) = E(LY - \theta)'(LY - \theta), \ (\theta,V) \in \mathcal{P}.$$

LY is called admissible among linear estimators if no other linear estimator is better than LY within model with parameter space $\mathcal{P}$.

The aim of this paper is to establish necessary and sufficient conditions for LY to be admissible under the above quadratic loss function when different sets of restrictions are placed on $\mathcal{V}$.

Throughout the paper we shall use the following notation:

n.n.d. = non-negative definite (symmetric).

p.d. = positive definite (symmetric).

p.s. = parameter space.

$[V]$ = the closed convex cone generated by n.n.d. matrix V.

$\mathcal{T}$ = $\{(\theta\theta',V):(\theta,V) \in \mathcal{P}\}$.

$[\mathcal{T}]$ = the smallest closed convex cone containing $\mathcal{T}$.

Note that $(\Phi,V) \in [\mathcal{T}]$ implies that Φ and V are n.n.d.

Also note that $(\Phi,0) \in [\mathcal{T}]$ for all n.n.d. matrices Φ.

Let $R(L|\Phi,V) = \mathrm{tr}[L'LV + (L-I)'(L-I)\Phi]$ for $(\Phi,V) \in [\mathcal{T}]$.

Note that $R(L|\theta,V) = R(L|\theta\theta',V)$ for $(\theta,V) \in \mathcal{P}$.

$\mathcal{R}(A)$ = linear manifold generated by the columns of matrix A.

$\mathcal{N}(A)$ = $[\mathcal{R}(A)]^{\perp}$ - orthogonal complement of $\mathcal{R}(A)$.

A' = the transpose of A.

A^+ = the Moore-Penrose inverse of A.

I = the unit matrix.

tr A = the trace of A.

$A \geq B$ denotes $A - B$ is n.n.d.

$A > B$ denotes $A - B$ is p.d.

$\mathcal{S}_{a\times b}$ = space of all $a \times b$ matrices.

$\mathcal{S}_{m\times n} \times \mathcal{S}_{a\times b}$ = product space.

$\mathcal{S}^2_{a\times b}$ = $\mathcal{S}_{a\times b} \times \mathcal{S}_{a\times b}$.

$[A,B]$ = tr AB'

$\mathcal{V}_n$ = the set of all n.n.d. matrices of order $n \times n$.

II. SOME USEFUL LEMMAS

Lemma 2.1.

Let V be an n.n.d. matrix of order $n \times n$ and of rank $r \geq 1$. Let A be any symmetric matrix such that AV is symmetric. Then A has r eigenvectors, say, $P_1,\ldots,P_r$, such that

$$V = \sum_{i=1}^{r} \tau_i P_i P_i' ,$$

where $\tau_i > 0$, $i=1,\ldots,r$.

Proof.

(i) First suppose that V is p.d. If AV is symmetric, then $W = V^{-\frac{1}{2}} AV^{\frac{1}{2}}$ is also symmetric. In fact,

$$W' = V^{\frac{1}{2}} A' V^{-\frac{1}{2}} = V^{-\frac{1}{2}} VA'V^{-\frac{1}{2}} = V^{-\frac{1}{2}} (AV)'V^{-\frac{1}{2}} = V^{-\frac{1}{2}} AVV^{-\frac{1}{2}} = V^{-\frac{1}{2}} AV^{\frac{1}{2}} = W.$$

Let $\lambda_1,\ldots, \lambda_n$ be the eigenvalues of W, and let $Q_1,\ldots,Q_n$ be normalized orthogonal eigenvectors of W with eigenvalues $\lambda_1,\ldots, \lambda_n$, respectively. Thus $WQ_i = \lambda_i Q_i$, $i=1,\ldots,n$. For $i=1,\ldots,n$ let

$$P_i = V^{\frac{1}{2}} Q_i .$$

Note that $A = V^{\frac{1}{2}} WV^{-\frac{1}{2}}$ and $Q_i = V^{-\frac{1}{2}} P_i$. With this notation, for $i=1,\ldots,n$

$$AP_i = AV^{\frac{1}{2}} Q_i = V^{\frac{1}{2}} WV^{-\frac{1}{2}} V^{\frac{1}{2}} Q_i = V^{\frac{1}{2}} WQ_i$$

$$= V^{\frac{1}{2}} (\lambda_i Q_i) = \lambda_i V^{\frac{1}{2}} V^{-\frac{1}{2}} P_i = \lambda_i P_i .$$

Thus P_i is an eigenvector of A with eigenvalue λ_i, $i=1,\ldots,n$. On the other hand, because $Q_1Q_1' + \ldots + Q_nQ_n' = I$ by assumption, we have

$$(V^{-\frac{1}{2}} P_1)(V^{-\frac{1}{2}} P_1)' + \ldots + (V^{-\frac{1}{2}} P_n)(V^{-\frac{1}{2}} P_n)' = I.$$

Hence

$$V^{-\frac{1}{2}} P_1 P_1' V^{-\frac{1}{2}} + \ldots + V^{-\frac{1}{2}} P_n P_n' V^{-\frac{1}{2}} = I,$$

which implies that $P_1P_1' + \ldots + P_nP_n' = V$, $i = 1,2,\ldots,n$ or

$$V = \tau_1(\tau_1^{-\frac{1}{2}} P_1)(\tau_1^{-\frac{1}{2}} P_1)' + \ldots + \tau_n(\tau_n^{-\frac{1}{2}} P_n)(\tau_n^{-\frac{1}{2}} P_n)' ,$$

where $\tau_i = P_i' P_i = Q_i' VQ_i > 0$, $i=1,\ldots,n$.

(ii) In case V is singular, there exists an orthogonal matrix R such that $RVR' = \mathrm{diag}(\sigma_1^2,\ldots,\sigma_r^2, 0,\ldots,0)$, where $\sigma_i^2 > 0$, $i=1,\ldots,r$, $r < n$. Observe that AV is symmetric if and only if $(RAR')(RVR')$ is symmetric. In fact $[(RAR')(RVR')]' = (R\,AV\,R')' = R(AV)'R' = R\,AV\,R' = (RAR')(RVR')$ if and only if $(AV)' = AV$. Thus without loss of generality we may assume that

$$V = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

where D is an $r \times r$ matrix, $1 \leq r < n$, and D is p.d. Partitioning correspondingly A leads to

$$AV = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_{11}D & 0 \\ A_{21}D & 0 \end{pmatrix}.$$

Because AV is symmetric by assumption, therefore $A_{21}D$ must be the $(n-r)\mathrm{x}r$ zero matrix and $A_{11}D$ must be a symmetric matrix. Because D is p.d., we conclude that $A_{21} = 0$. Moreover, because $A_{11}D$ is symmetric, it follows from part (i) that there exist normalized eigenvectors $S_1,\ldots,S_r$ of A_{11} such that $D = \tau_1 S_1 S_1 + \ldots + \tau_r S_r S_r$, $\tau_i > 0$, where $A_{11}S_i = \lambda_i S_i$, $i=1,\ldots,r$.

Letting $$P_i = \begin{pmatrix} S_i \\ 0 \end{pmatrix} \in \mathcal{R}^n$$

we note that

$$AP_i = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} S_i \\ 0 \end{pmatrix} = \begin{pmatrix} A_{11}S_i \\ A_{21}S_i \end{pmatrix}.$$

Because $A_{11}S_i = \lambda_i S_i$ and because $A_{21} = 0$, it follows that

$$AP_i = \begin{pmatrix} \lambda_i S_i \\ 0 \end{pmatrix} = \lambda_i P_i.$$

Thus $P_1,\ldots,P_r$ are real eigenvectors of matrix A. Moreover,

$$\tau_1 P_1 P_1' + \ldots + \tau_r P_r P_r' = \tau_1 \begin{pmatrix} S_1 \\ 0 \end{pmatrix} (S_1',0) + \ldots + \tau_r \begin{pmatrix} S_r \\ 0 \end{pmatrix} (S_r',0)$$

$$= \tau_1 \begin{pmatrix} S_1 S_1' & 0 \\ 0 & 0 \end{pmatrix} + \ldots + \tau_r \begin{pmatrix} S_r S_r' & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \tau_1 S_1 S_1' + \ldots + \tau_r S_r S_r' & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = V.$$

This terminates the proof.

Lemma 2.2.

If matrix

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix}$$

is n.n.d., then $\mathcal{R}(V_{12}') \subset \mathcal{R}(V_{22})$.

Proof.

This is a known result.

Lemma 2.3.

Let V_{12} be an $r \times (n-r)$ matrix and let V_{22} be an $(n-r)\times(n-r)$ n.n.d. matrix. where $1 \leqslant r < n$. If $R(V'_{12}) \subset R(V_{22})$, then there exists an n.n.d. matrix V_{11} of order $r \times r$ such that

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V'_{12} & V_{22} \end{pmatrix}$$

is n.n.d.

Proof.

This is a known result.

Lemma 2.4.

Let V and Φ be symmetric. If $L(V + \Phi) = \Phi$, then $(LV)' = LV$ and $(L\Phi)' = L\Phi$.

Proof.

Because $\Phi = L(V + \Phi)$ and because Φ is symmetric,

$$\Phi = (V + \Phi)L'.$$

Hence

$$L\Phi = L(V + \Phi)L',$$

which is a symmetric matrix. Moreover,

$$LV = \Phi - L\Phi.$$

Hence

$$(LV)' = \Phi - (L\Phi)' = \Phi - L\Phi = LV.$$

Thus LV is symmetric as asserted.

Lemma 2.5.

Let V and Φ be n.n.d. matrices. If $L(V + \Phi) = \Phi$ and if $V + \Phi$ is p.d., then the eigenvalues of L are in the interval $[0,1]$. Further, if V is p.d. the eigenvalues of L are in the interval $[0,1)$.

Proof.

This is a known result.

Lemma 2.6.

Let X and V be $n \times n$ matrices. Let ψ: $S_{n\times n} \to R$ be defined by $\psi(X) = \operatorname{tr} XVX'$. Then ψ is convex if V is n.n.d. and strictly convex if V is p.d.

Proof.

Letting $X = (x_1,\dots,x_n)$, where $x_i \in \mathcal{R}$, we find that $\psi(X) = x_1'Vx_1 + \dots + x_n'Vx_n$. Because the sum of convex (strictly convex) functions is a convex (strictly convex), the assertion follows.

Lemma 2.7.

Let $\mathcal{R}(A) \subset \mathcal{R}(V)$, where V is a n.n.d. matrix. If $\operatorname{tr} A'VA = 0$ then $A = 0$.

Proof.

As in the proof of Lemma 2.6 we write $\operatorname{tr} A'VA = a_1'Va_1 + \dots + a_n'Va_n$, where $A = (a_1,\dots,a_n)$. From the assumptions it now follows that $a_i'Va_i = 0$ and that $a_i = V^{\frac{1}{2}}b_i$, where $b_i \in \mathcal{R}$, for $i = 1,\dots,n$. Combining these two formulas leads to $Vb_i = 0$, which implies that $a_i = 0$ for $i = 1,\dots,n$. Hence $A = 0$.

Lemma 2.8.

Let R be an $n \times n$ orthogonal matrix. If LY is admissible within model p.s. $\mathcal{R}^n \times \mathcal{V}$, then MZ, where $M = RLR'$, is admissible within model with p.s. $\mathcal{R}^n \times \{RVR' : V \in \mathcal{V}\}$. Here Z stands for an n-variate random vector with p.s. $\mathcal{R}^n \times \{RVR' : V \in \mathcal{V}\}$.

Proof.

See the paper of A. Cohen [1].

III. LA MOTTE'S THEOREM AND ITS EXTENSION

In 1978, R. La Motte [2] announced a theorem that establishes a tractable characterization of admissible estimators. In our considerations we shall need only the part of La Motte's theorem which states that if LY is admissible within model with p.s. $\mathcal{P} = R^n \times \mathcal{V}$, then there exists a nonzero pair $(\Phi,V) \in [\mathcal{T}]$ such that $L(\Phi + V) = \Phi$. Note that this condition determines L uniquely when $\Phi + V$ is p.d. A drawback of La Motte's theorem is that $L(\Phi + V) = \Phi$ may reduce to $L\Phi = \Phi$, Φ singular, even when all nonzero matrices in $\mathcal{V}$ are p.d. If this is the case, then L is not uniquely determined by $L\Phi = \Phi$. To overcome this drawback, we may use Theorem 3.4 established in this Section. This theorem applies only to matrices L which have a particular structure.

First we recall a result from La Motte's paper [2] which gives necessary and sufficient conditions for estimator LY to minimize $R(\cdot|\Phi,V)$. These conditions may be already found in a paper of C.R. Rao [4] in the particular case where V is p.d. and $\Phi = \theta\theta'$.

Theorem 3.1.

Let $(\Phi,V) \in [\mathcal{T}]$. Then

(i) there exists a $n \times n$ matrix L such that $L(V + \Phi) = \Phi$;

(ii) L minimizes $R(\cdot|\Phi,V)$ if an only if $L(V + \Phi) = \Phi$;

(iii) there exists a unique matrix minimizing $R(\cdot|\Phi,V)$ if and only if $\Phi + V$ is p.d.

Proof.

If $(\Phi,V) \in [T]$, then both Φ and V are n.n.d. In such a case $V + \Phi$ is p.d. and, therefore

$$R(\Phi) \subset R(V + \Phi).$$

Consequently, there exists a matrix L such that $L(V + \Phi) = \Phi$.

To prove assertion (ii), recall that

$$R(L|\Phi,V) = \operatorname{tr} L'LV + \operatorname{tr} (I - L)'(I - L)\Phi.$$

Since $(V + \Phi)(V + \Phi)^{+}$ is a projection operator on $R(V + \Phi)$, we have $(V + \Phi)(V + \Phi)^{+} = \Phi$. Hence

$$\begin{aligned} R(L|\Phi,V) &= \operatorname{tr} L'LV + \operatorname{tr} [\Phi - L\Phi - L'\Phi + L'L\Phi] \\ &= \operatorname{tr} L'L(V + \Phi) + \operatorname{tr} [\Phi - L\Phi - L'\Phi] \\ &= \operatorname{tr} L(V + \Phi)(V + \Phi)^{+}(V + \Phi)L' - \operatorname{tr} [L(V + \Phi)(V + \Phi)^{+}\Phi \\ &\qquad + L'\Phi(V + \Phi)(V + \Phi)^{+} - \Phi] \\ &= \operatorname{tr} [L(V + \Phi) - \Phi](V + \Phi)^{+}[L(V + \Phi) - \Phi]' + \operatorname{tr} [\Phi - \\ &\qquad \Phi(V + \Phi)^{+}\Phi]. \end{aligned}$$

We have used here the fact that

$$\begin{aligned} \operatorname{tr} \Phi(V + \Phi)^{+}[L(V + \Phi)]' &= \operatorname{tr} \Phi(V + \Phi)^{+}(V + \Phi)L' \\ &= \operatorname{tr} L'\Phi(V + \Phi)^{+}(V + \Phi) \\ &= \operatorname{tr} L'(V + \Phi)(V + \Phi)^{+} \Phi \\ &= \operatorname{tr} L'\Phi. \end{aligned}$$

Hence

$$R(L|\Phi,V) = \operatorname{tr}[L(V+\Phi) - \Phi](V+\Phi)^{+}[L(V+\Phi) - \Phi]' + \operatorname{tr}[\Phi - \Phi(V+\Phi)^{+}\Phi].$$

Note that L does not appear in the second term on the right hand side.

Because $(V + \Phi)^{+}$ is n.n.d., the first term is nonnegative. Hence L minimizes $R(\cdot|\Phi,V)$ if and only if $L(V + \Phi) = \Phi$ by Lemma 2.7.

Assertion (iii) is obvious. In fact, if $V + \Phi$ is p.d., then L is uniquely determined by $L = \Phi(V + \Phi)^{-1}$.

The following two simple corollaries to Theorem 3.1 are very useful in further considerations.

Corollary 3.1.

Let L and M be $n \times n$ matrices. If $LP = \lambda P$, where $\lambda \in [0,1]$ and if $R(M|\theta,V) \leq R(L|\theta,V)$ for $V = (1 - \lambda)PP'$ and $\Phi = \lambda PP'$, then $MP = \lambda P$.

Proof.

If $\Phi = \lambda PP'$ and if $V = (1 - \lambda)PP'$, then $\Phi + V = PP'$. Consequently, $L(\Phi + V) = (LP)P' = (\lambda P)P' = \Phi$. Thus L minimizes $R(\cdot|\Phi,V)$ at $(\Phi,V) = (\lambda PP',(1 - \lambda)PP')$. Because $R(M|\Phi,V) \leq R(L|\Phi,V)$ we conclude that $R(M|\Phi,V) = R(L|\Phi,V)$ and therefore $M(\Phi + V) = \Phi$. Using $\Phi + V = PP'$, we note that $M(\Phi + V) = MPP' = \lambda PP'$. Thus $MP = \lambda P$ as asserted.

Corollary 3.2.

If MY is as good as LY and if $LP = P$, then $MP = P$.

Proof.

By assumption $R(M|\Phi,V) \leq R(L|\Phi,V)$ for $V = 0$ and $\Phi = PP'$. Now $LP = P$ yields $LPP' = PP'$. Hence $MPP' = PP'$ by Theorem 3.1. Consequently, $MP = P$.

Theorem (La Motte) 3.2.

If LY is admissible within model with p.s. $P = R^n \times V$, then there exists a nonzero point $(\Phi,V) \in [T]$ such that L minimizes $R(\cdot|\Phi,V)$.

Proof.

As already shown

$$R(L|\Phi,V) = \mathrm{tr}\,[L(\Phi + V) - \Phi](\Phi + V)^{+}[L(\Phi + V) - \Phi] + \mathrm{tr}\,[\Phi - \Phi(\Phi + V)\Phi].$$

Let

$$E = \{(V,\Phi):\ V \text{ and } \Phi \text{ are n.n.d., } \mathrm{tr}\,(VV + \Phi\Phi) = 1\}.$$

Let W be the convex hull of $[T] \cap E$. It should be noted that $(0,0) \notin W$. Suppose to the contrary that $(0,0) \in W$. Then there would exist points $(\Phi_1,V_1),\ldots,(\Phi_s,V_s)$ in $[T] \cap E$ such that $\sum_{i=1}^{s} \tau_i \Phi_i = 0$

and $\sum_{i=1}^{s} \tau_i V_i = 0$ where the τ's are positive. However, since the V's and Φ's are n.n.d., therefore $V_1 = \ldots = V_s = 0$ and $\Phi_1 = \ldots = \Phi_s = 0$. This contradicts the assumption that $(\Phi_i, V_i) \in E$ for $i = 1,\ldots,s$.

It should also be noted that for each $(\Phi,V) \in [T]$ there exists a point $(\Phi_0,V_0) \in \mathcal{W}$ such that $(\Phi,V) = \tau(\Phi_0,V_0)$, where $\tau \in R^+$.

Now define a function

$$f\colon\ \mathcal{W} \to S_{n\times n}$$

by

$$f(\Phi,V) = L(\Phi + V) - \Phi.$$

This is a linear mapping and hence preserves compactness and convexity. Thus the set $W = \{f(\Phi,V)\colon\ (\Phi,V) \in \mathcal{W}\}$ is compact and convex in the product space $S^2_{n\times n}$. We have to show that if LY is admissible, then there exists a nonzero point $(\Phi,V) \in [T]$, such that L minimizes $R(\cdot|\Phi,V)$. To prove this, it is sufficient to show that $(0,0) \in W$. In fact, if $(0,0) \notin W$ then there exists a pair $(\Phi,V) \in \mathcal{W}$ such that $L(V + \Phi) - \Phi = 0$, and, therefore, L minimizes $R(\cdot|\Phi,V)$ by Theorem 3.1.

Suppose to the contrary that $(0,0) \notin W$. Because W is closed and convex, the separating hyperplane theorem assures the existence of a matrix H in $S_{n\times n}$ such that for all $X \in W$

$$\text{(3.1)} \qquad \operatorname{tr} XH' < 0.$$

Now for $\gamma \in R$ define matrix $M = L + \gamma H$. Note that

$$\begin{aligned} R(M|\Phi,V) &= \operatorname{tr} (L + \gamma H)'(L + \gamma H)V + \operatorname{tr} [(I - L - \gamma H)'(I - L - \gamma H)\Phi] \\ &= \operatorname{tr} L'LV + \gamma \operatorname{tr} [L'HV + \gamma \operatorname{tr} H'LV + \gamma^2 \operatorname{tr} H'HV \\ &\qquad + \operatorname{tr} (I - L)'(I - L)\Phi - \gamma \operatorname{tr} (I - L)'H\Phi \\ &\qquad - \gamma \operatorname{tr} H'(I - L) + \gamma^2 \operatorname{tr} H'H\Phi \\ &= \operatorname{tr} L'LV + \operatorname{tr} (I - L)'(I - L)\Phi + 2\gamma \operatorname{tr} [L'HV - (I - L)'H\Phi] \\ &\qquad + \gamma^2 \operatorname{tr} H\Phi H'. \end{aligned}$$

Hence

$$R(M|\Phi,V) - R(L|\Phi,V) = \gamma^2 \operatorname{tr} H\Phi H' + 2\gamma \operatorname{tr} [L(\Phi + V) - \Phi]H'.$$

Because

$$\operatorname{tr} [L(V + \Phi) - \Phi]H' < 0$$

for $(\Phi,V) \in \mathcal{W}$, we conclude that

$$\operatorname{tr} H(V + \Phi)H' > 0$$

for all $(\Phi,V) \in \mathcal{W}$. In fact, because $V + \Phi$ is n.n.d., $\operatorname{tr} H(V + \Phi)H' =$ $\operatorname{tr} H(V + \Phi)^{\frac{1}{2}}H(V + \Phi)^{\frac{1}{2}}$. This shows that $\operatorname{tr} H(V + \Phi)H' \geq 0$. Now suppose that $\operatorname{tr} H(V + \Phi)H' = 0$. Then $H(V + \Phi)^{\frac{1}{2}} = 0$, and, consequently, $(V + \Phi)H' = 0$. On the other hand, for all $(\Phi,V) \in \mathcal{W}$ we have $\operatorname{tr} [L(V + \Phi) - \Phi]H' = \operatorname{tr} [L(V + \Phi) - (V + \Phi)^+(V + \Phi)H' =$ $\operatorname{tr} [L - \Phi(V + \Phi)^+](V + \Phi)H' = 0$ which contradicts (3.1).

Thus, letting $\pi(\Phi,V,\gamma) = R(M|\Phi,V) - R(L|\Phi,V)$,

$$\pi(\Phi,V,\gamma) = a\gamma^2 + b\gamma,$$

where

$$a = a(\Phi,V) = \mathrm{tr}\, H(V + \Phi)H' > 0,$$

while

$$b = b(\Phi,V) = \mathrm{tr}\, [L(\Phi + V) - \Phi]H' < 0$$

for all $(\Phi,V) \in \mathcal{W}$.

Then for arbitrary, but fixed $(\Phi,V) \in \mathcal{W}$, the quadratic polynomial $\pi(\Phi,V,\gamma)$ in γ achieves its minimum value of $-b^2/a$ when $\gamma = g = -b^2/a$. Since g, considered as a mapping from $\mathcal{W}$ to $\mathcal{R}$ is continuous and strictly positive on the compact set $\mathcal{W}$, there exists $\varepsilon > 0$ such that $g(\Phi,V) \geqslant \varepsilon$ for all $(\Phi,V) \in \mathcal{W}$. We see that

$$\varepsilon^{-1}\pi(\Phi,\varepsilon) \leqslant g(\Phi,V)^{-1}\pi(\Phi,V,g(\Phi,V)) < 0$$

for all $(\Phi,V) \in W$. Hence

$$R(L + \varepsilon H|\Phi,V) \leqslant R(L|\Phi,V)$$

for all $(\Phi,V) \in [\mathcal{T}]$ and strict inequality if $(V_{12},V_{22},\Phi_{22})$ is in $\mathcal{W}$. This contradicts the assumption that LY is admissible and shows that $(0,0) \in W$.

Now we proceed to establish the above mentioned extension of La Motte's theorem. This extension refers only to the subset L of all $n \times n$ matrices of the form

$$L = \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{pmatrix},$$

where L_{11} is an arbitrary, fixed $r \times r$ matrix, while $1 \leqslant r < n$. First we show some preliminary results.

Lemma 3.1.

Let

$$L = \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{pmatrix}, \; V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix} \geqslant 0 \text{ and } \Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}' & \Phi_{22} \end{pmatrix} \geqslant 0,$$

where L_{11}, V_{11} and Φ_{11} are $r \times r$ matrices. Then

$$R(L|\Phi,V) = R(L_{22}|\Phi_{22},V_{22}) + \operatorname{tr}\left[(I,L_{12})A\begin{pmatrix} I \\ L_{12}' \end{pmatrix} + (I,-L_{12})B\begin{pmatrix} I \\ -L_{12} \end{pmatrix}\right],$$

where

$$A = \begin{pmatrix} L_{11}V_{11}L_{1}' & L_{11}V_{12} \\ V_{12}'L_{11}' & V_{22} \end{pmatrix},$$

while

$$B = \begin{bmatrix} (I - L_{11})\Phi_{11}(I - L_{11})' & (I - L_{11})\Phi_{12} \\ \Phi_{12}'(I - L_{11})' & \Phi_{22} \end{bmatrix}.$$

Proof.

First note that

$$\Delta_1 = (I,L_{12})A\begin{bmatrix} I \\ L_{12}' \end{bmatrix} = (I,L_{12})\begin{bmatrix} L_{11} & 0 \\ 0 & I \end{bmatrix} V \begin{bmatrix} L_{11}' & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} I \\ L_{12}' \end{bmatrix}$$

$$= (L_{11},L_{12})V\begin{bmatrix} L_{11}' \\ L_{12}' \end{bmatrix}$$

so that

$$\operatorname{tr} \Delta_1 = \operatorname{tr} \begin{bmatrix} L_{11}' \\ L_{12}' \end{bmatrix} (L_{11},L_{12})V.$$

Similarly

$$\Delta_2 = (I,-L_{12})B\begin{bmatrix} I \\ -L_{12}' \end{bmatrix} = (I - L_{11},-L_{12})\Phi\begin{bmatrix} I-L_{11}' \\ -L_{12}' \end{bmatrix}$$

and

$$\operatorname{tr} \Delta_2 = \operatorname{tr} \begin{bmatrix} I-L_{11}' \\ -L_{12}' \end{bmatrix}(I - L_{11},-L_{12})\Phi.$$

Next observe that

$$R(L|\phi,V) = \mathrm{tr}\begin{bmatrix} L_{11}' & 0 \\ L_{12}' & L_{22}' \end{bmatrix}\begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix} V$$

$$+ \mathrm{tr}\begin{bmatrix} I - L_{11}' & 0 \\ -L_{12}' & I - L_{22}' \end{bmatrix}\begin{bmatrix} I - L_{11} & -L_{12} \\ 0 & I - L_{22} \end{bmatrix}\Phi$$

$$= \mathrm{tr}\begin{bmatrix} L_{11}'L_{11} & L_{11}'L_{12} \\ L_{12}'L_{11} & L_{12}'L_{12} + L_{22}'L_{22} \end{bmatrix} V$$

$$+ \mathrm{tr}\begin{bmatrix} (I - L_{11}')(I - L_{11}) & -(I - L_{11}')L_{12} \\ -L_{12}'(I - L_{11}) & L_{12}'L_{12} + (I - L_{22}')(I - L_{22}) \end{bmatrix}\Phi$$

$$= \mathrm{tr}\,\Delta_1 + \mathrm{tr}\begin{bmatrix} 0 & 0 \\ 0 & L_{22}'L_{22} \end{bmatrix} V + \mathrm{tr}\,\Delta_2$$

$$+ \mathrm{tr}\begin{bmatrix} 0 & 0 \\ 0 & (I - L_{22}')(I - L_{22}) \end{bmatrix}\Phi,$$

as asserted.

Theorem 3.3.

Let L_{11} be an arbitrary fixed $r \times r$ matrix, $1 \leqslant r < n$. Matrix

$L = \begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix} \in \mathcal{L}$ minimizes $R(\cdot|\Phi,V)$ among $\mathcal{L}$ if and only if

$$(3.2)\qquad \begin{aligned} L_{12}(V_{22} + \Phi_{22}) &= -L_{11}V_{12} + (I - L_{11})\Phi_{12} \\ L_{22}(V_{22} + \Phi_{22}) &= \Phi_{22}. \end{aligned}$$

Proof.

Let A and B be defined as in Lemma 3.1. Put

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}' & B_{22} \end{bmatrix},$$

where A_{11} and B_{11} are $r \times r$ matrices. Because A and B are n.n.d., therefore

$$\mathcal{R}(A_{12}') \subset \mathcal{R}(A_{22} + B_{22})$$

and

$$\mathcal{R}(B_{12}') \subset \mathcal{R}(A_{22} + B_{22}).$$

Hence there exists a matrix H such that

$$A_{12}' - B_{12}' = (A_{22} + B_{22})H.$$

For convenience let $X = L_{12}$. Note that by Lemma 3.1

$$R(L|\Phi,V) = R(L_{22}|\Phi_{22},V_{22}) + \text{tr}\ [(I,X)A\begin{pmatrix} I \\ X' \end{pmatrix} + (I,-X)B\begin{pmatrix} I \\ X' \end{pmatrix}].$$

Also note that

$$(I,X)A\begin{pmatrix} I \\ X' \end{pmatrix} + (I,-X)B\begin{pmatrix} I \\ X' \end{pmatrix} = A_{11} + X(A'_{12} - B'_{12}) + (A_{12} - B_{12})X'$$
$$+ X(A_{22} + B_{22})X'$$

and that

$$[X(A_{22} + B_{22}) + A_{12} - B_{12}](A_{22} + B_{22}) + [X(A_{22} + B_{22}) - A_{12} - B_{12}]'$$
$$= X(A'_{12} - B'_{12}) + (A_{12} - B_{12})X' + X(A_{22} + B_{22})X'$$
$$+ (A_{12} - B_{12})(A_{22} + B_{22}) + (A_{12} - B_{12})'.$$

Hence

$$R(L|\Phi,V) = R(L_{22}|\Phi_{22},V_{22}) + [X(A_{22} + B_{22}) + A_{12} - B_{12}](A_{22} + B_{22})$$
$$+ [X(A_{22} + B_{22}) + A_{12} - B_{12}]$$
$$+ A_{11} - (A_{12} - B_{12})(A_{22} + B_{22})^{+}(A_{12} - B_{12})'.$$

Because $(A_{22} + B_{22})^{+}$ is n.n.d. and because

$$\mathcal{R}((A_{22} + B_{22})X + A_{12}' - B_{12}') \subset \mathcal{R}(A_{22} + B_{22})$$

it follows from Lemma 2.7 and Theorem 3.1 that L minimizes $R(\cdot|\Phi,V)$ among $\mathcal{L}$ if and only if conditions (3.2) are satisfied.

Let us now introduce the following notation. As above, let $\mathcal{P} = \mathcal{R}^n \times \mathcal{V}$ be the considered parameter space. Let $1 \leq r < n$ and let V and E be subsets of $S_{r\times(n-r)} \times S_{(n-r)\times(n-r)}$ defined by $V = \{(V_{12},V_{22},\Phi_{22})$: there exist matrices V_{11}, Φ_{11}, Φ_{12} such that

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix}, \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}' & \Phi_{22} \end{pmatrix} \in [T]\}$$

and by

$$E = \{(V_{12},V_{22},\Phi_{22}): \ \mathrm{tr}\, V_{12}V_{12}' + \mathrm{tr}\, V_{22}V_{22}' + \mathrm{tr}\, \Phi_{22}\Phi_{22} = 1\},$$

respectively. Let W be the convex hull of $V \cap E$. Note that $(0_{r\times(n-r)}, 0_{(n-r)\times(n-r)}, 0_{(n-r)\times(n-r)}) \notin W$.

Theorem 3.4.

Let W be compact. If $L = \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{pmatrix}$ and if LY is admissible, then there exists a point $(\Phi,V) \in [T]$ with $(V_{12},V_{22},\Phi_{22})$ in W such

that L minimizes $R(\cdot|\Phi,V)$ among the class L. Moreover, (Φ,V) may be selected in such a way that L_{12} and L_{22} meet conditions (3.2), i.e.,

$$L_{12}(V_{22} + \Phi_{22}) = -L_{11}V_{12} + (I - L_{11})\Phi_{12}$$

$$L_{22}(V_{22} + \Phi_{22}) = \Phi_{22}.$$

Proof.

Theorem 3.4 is proved essentially in the same way as Theorem 3.2, although the details are more complex. Let $S = S_{r\times(n-r)} \times S_{(n-r)\times(n-r)}$. Define a function

$$f: \mathcal{W} \to S$$

by

$$f(V_{12},V_{22},\Phi_{22}) = [L_{12}(V_{22} + \Phi_{22}) + L_{12}V_{12} - (I - L_{11})\Phi_{12}, L_{22}(V_{22} + \Phi_{22}) - \Phi_{22}.$$

Let

$$W = \{f(V_{12},V_{22},\Phi_{22}): (V_{12},V_{22},\Phi_{22}) \in \mathcal{W}\}.$$

Note that f is a linear mapping. Therefore W is compact and convex, because $\mathcal{W}$ is compact and convex by assumption. To prove the theorem, it is sufficient to show that $(0_{r\times(n-r)}, 0_{(n-r)\times(n-r)}) \in W$. Suppose that $(0,0) \notin W$. Then the separating hyperplane theorem assures the existence of

a pair of matrices (H_{12},H_{22}) in T such that for all $(X_1,X_2) \in W$

(3.3) $$\operatorname{tr} X_1H_1' + \operatorname{tr} X_2H_2' < 0.$$

Now for $\gamma \in R$ define matrix

$$M = L + \gamma\begin{pmatrix} 0 & H_{12} \\ 0 & H_{22} \end{pmatrix} = \begin{pmatrix} L_{11} & L_{12}+\gamma H_{12} \\ 0 & L_{22}+\gamma H_{22} \end{pmatrix}.$$

Clearly, $M \in L$.

Let A and B be defined as in Lemma 3.1. Then

$$R(L|\Phi,V) = R(L_{22}|\Phi_{22},V_{22}) + \operatorname{tr}\ [(I,L_{12})A\begin{pmatrix} I \\ L_{12}' \end{pmatrix} + (I,-L_{12})B\begin{pmatrix} I \\ -L_{12}' \end{pmatrix}]$$

and

$$R(M|\Phi,V) = R(L_{22}+\gamma H_{22}|\Phi_{22},V_{22}) + \operatorname{tr}\ [(I,L_{12}+\gamma H_{12})A\begin{pmatrix} I \\ L_{12}'+\gamma H_{12}' \end{pmatrix}$$

$$+ (I,-L_{12}-\gamma H_{12})B\begin{pmatrix} I \\ -L_{12}'-\gamma H_{12}' \end{pmatrix}].$$

We need to find

$$\pi(\Phi,V,\gamma) = R(M|\Phi,V) - R(L|\Phi,V).$$

First we find

$$R(L_{22}+\gamma H_{22}|\Phi_{22},V_{22}) - R(L_{22}|\Phi_{22},V_{22}) = \gamma^2 \operatorname{tr} H_{22}(V_{22}+\Phi_{22})H_{22}'$$

$$+ 2\gamma \operatorname{tr}\ [L_{22}(V_{22}+\Phi_{22}) - \Phi_{22}]H_{22}'.$$

The remaining term becomes

$$\gamma^2 \operatorname{tr} H_{12}(A_{22} + B_{22})H_{12}' + 2\gamma \operatorname{tr} [(A_{12} - B_{12} + L_{12}(A_{22} + B_{22})]H_{12}'.$$

Thus

$$\begin{aligned}\pi(\Phi,V,\gamma) &= \gamma^2 \operatorname{tr} [H_{22}(V_{22} + \Phi_{22})H_{22}' + H_{12}(V_{22} + \Phi_{22})H_{12}'] \\ &\qquad + 2\gamma \operatorname{tr} \{[L_{22}(\Phi_{22} + V_{22}) - \Phi_{22}]H_{22}' \\ &\qquad + [L_{12}(V_{22} + \Phi_{22}) + L_{11}V_{12} - (I - L_{11})\Phi_{12}]H_{12}'\} \\ &= a\gamma^2 + 2b\gamma.\end{aligned}$$

Now observe that $a > 0$ and $b < 0$ for all $(V_{12},V_{22},\Phi_{22}) \in \mathcal{W}$. Inequality $a \leqslant 0$ leads to a contradiction with (3.3). We may now proceed along the same lines as in the proof of Theorem 3.2, and conclude that MY is better than L^0Y. This contradicts the assumption that LY is admissible. Hence $(0,0) \in W$, which yields the desired result by a similar argument as in the proof of Theorem 3.2.

In case $L_{11} = I$ we have the following two corollaries.

Corollary 3.3.

Suppose that $L = \begin{bmatrix} I & L_{12} \\ 0 & L_{22} \end{bmatrix}$, where I is $r \times r$, and that the corresponding set $\mathcal{W}$ is compact. If LY is admissible then there exists a nonzero point $(\Phi,V) \in [T]$ with $(V_{12},V_{22},\Phi_{22}) \in \mathcal{W}$ such that L minimizes

$R(\cdot|\Phi,V)$ among the class of matrices of the form $\begin{bmatrix} I & L_{12} \\ 0 & L_{22} \end{bmatrix}$. Moreover, (Φ,V) may be selected in such a way that L_{12} and L_{22} meet the following two conditions:

$$\begin{aligned} L_{12}(V_{22} + \Phi_{22}) &= -V_{12} \\ L_{22}(V_{22} + \Phi_{22}) &= \Phi_{22}. \end{aligned} \tag{3.4}$$

Corollary 3.4.

Suppose that $L = \begin{bmatrix} I & L_{12} \\ 0 & L_{22} \end{bmatrix}$, where I is $r \times r$, has an r-fold degeneracy for the eigenvalue $\lambda = 1$, i.e., that $L_{12}Q \neq 0$ for all $Q \in R^{n-r}$ such that $L_{22}Q = Q$. If L meets conditions (3.4) at a nonzero point $(V_{12},V_{22},\Phi_{22})$, then V_{22} can not be the zero matrix.

Proof.

Suppose on the contrary that $V_{22} = 0$. Because $V_{22} = 0$ implies that $V_{12} = 0$, conditions (3.4) reduce to

$$\begin{aligned} L_{12}\Phi_{22} &= 0 \\ L_{22}\Phi_{22} &= \Phi_{22}. \end{aligned}$$

This shows that there exists a nonzero vector $Q \in R^{n-r}$ such that $L_{22}Q = Q$, $L_{12}Q = 0$, because $\Phi_{22} \neq 0$ by assumption. This contradicts

the assumption that L has an r-fold degeneracy for eigenvalue $\lambda = 1$. Thus $V_{22} \neq 0$.

A. Cohen has shown in [1] that LY is admissible within model p.s. $R^n \times [I]$ if and only if L is symmetric and the eigenvalues of L are in the closed interval $[0,1]$. The main element in his proof was the observation that MY, where $M = [I - (I - L)'(I - L)]^{\frac{1}{2}}$ is better than LY if L is asymmetric. Using a lemma due to N. Shinozaki, C.R. Rao [4] extended this result to models with p.s. $R^n \times [V]$, where V may be any p.d. matrix. Now we shall show that Cohen's theorem and its extension follow straightforward from Theorems 3.2 and 3.4. In the last section we shall prove a theorem (see Theorem 5.2) which includes these results as a particular case.

Theorem 3.5.

Consider a model with p.s. $P = R^n \times [V]$, where V is a p.d. matrix. In order that LY is admissible within model with p.s. P it is necessary and sufficient that LV is symmetric and the eigenvalues of L are in the closed interval $[0,1]$.

Proof.

(i) To show that the conditions of Theorem 3.5 are sufficient, let us first assume that the eigenvalues of L are in $[0,1)$. Because matrix $V^{-\frac{1}{2}}LV^{\frac{1}{2}}$ is symmetric, it can be presented in the form

$$V^{-\frac{1}{2}}LV^{\frac{1}{2}} = \sum_{i=1}^{n} \lambda_i P_i P_i',$$

where $P_1,\ldots,P_n$ are linearly independent orthogonal eigenvalues of L corresponding to eigenvalues $\lambda_1,\ldots,\lambda_n$, respectively.

Hence

$$L = \sum_{i=1}^{n} \lambda_i V^{\frac{1}{2}} P_i P_i' V^{-\frac{1}{2}}.$$

Because by assumption $0 \leq \lambda_i < 1$, it follows that $0 < 1 - \lambda_i \leq 1$. Consequently,

$$(I - V^{-\frac{1}{2}} L V^{\frac{1}{2}})^{-1} = \sum_{i=1}^{n} \frac{1}{1 - \lambda_i} P_i P_i',$$

so that

$$(I - L)^{-1} = \sum_{i=1}^{n} \frac{1}{1 - \lambda_i} V^{\frac{1}{2}} P_i P_i' V^{-\frac{1}{2}}.$$

Hence

$$(I - L)^{-1} L = \sum_{i=1}^{n} \frac{\lambda_1}{1 - \lambda_i} V^{\frac{1}{2}} P_i P_i' V^{-\frac{1}{2}}$$

and, finally,

$$(I - L)^{-1} LV = \sum_{i=1}^{n} \frac{\lambda_i}{1 - \lambda_i} V^{\frac{1}{2}} P_i P_i' V^{\frac{1}{2}}.$$

We have shown that if LV is symmetric and if the eigenvectors of L are in $[0,1)$, then $\Phi = (I - L)^{-1} LV$ is n.n.d. Now rewrite $\Phi = (I - L)^{-1} LV$ in the form of $L(V + \Phi) = \Phi$. Theorem 3.1 implies that L minimizes $R(\cdot|\Phi,V)$, and since $V + \Phi$ is p.d., we conclude that LY is admissible.

(ii) If $\lambda = 1$ is an eigenvalue of L, then L may be represented in the form

$$L = \begin{pmatrix} I & L_{12} \\ 0 & L_{22} \end{pmatrix}, \quad I \text{ is } r \times r,$$

where r is the maximal number of linearly independent eigenvectors of L corresponding to $\lambda = 1$. Because L is semisimple, all the eigenvalues of L_{22} are in the interval $[0,1)$. Now observe that

$$LV = \begin{pmatrix} I & L_{12} \\ 0 & L_{22} \end{pmatrix} \begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix} = \begin{pmatrix} V_{11} + L_{12}V_{12}' & V_{12} + L_{12}V_{22} \\ L_{22}V_{12}' & L_{22}V_{22} \end{pmatrix}.$$

Since LV is symmetric, we obtain that

(α) $L_{22}V_{22}$ is symmetric

(β) $V_{12} + L_{12}V_{22} = (L_{22}V_{12}')' = V_{12}L_{22}'$

(γ) $L_{12}V_{12}'$ is symmetric.

As shown in part (i), it follows from (α) that $(I - L_{22})^{-1}L_{22}V_{22}$ is a n.n.d. matrix. Put $\Phi_{22} = (I - L_{22})^{-1}L_{22}V_{22}$. From ($\beta$) it follows that

$$L_{12} = V_{12}(L_{22}' - I)V_{22}^{-1}.$$

Using $\Phi_{22} = (I - L_{22})^{-1}L_{22}V_{22}$ we note that

$$
\begin{aligned}
L_{12}(V_{22}+\Phi_{22}) &= V_{12}(L_{22}'-I)V_{22}^{-1}[V_{22}+V_{22}L_{22}'(I-L_{22}')^{-1}] \\
&= V_{12}(L_{22}'-I)[I+L_{22}'(I-L_{22}')^{-1} \\
&= V_{12}(L_{22}'-I)[(I-L_{22}')+L_{22}'](I-L_{22}')^{-1} \\
&= V_{12}(L_{22}'-I)(I-L_{22}')^{-1} \\
&= -V_{12}.
\end{aligned}
$$

Consequently,

$$
L = \begin{bmatrix} I & L_{12} \\ 0 & L_{22} \end{bmatrix}
$$

meets conditions (3.2) appearing in Theorem 3.3, i.e.,

$$
\begin{aligned}
L_{12}(V_{22}+\Phi_{22}) &= -V_{12} \\
L_{22}(V_{22}+\Phi_{22}) &= \Phi_{22}.
\end{aligned}
$$

Hence the matrix $L = \begin{bmatrix} I & L_{12} \\ 0 & L_{22} \end{bmatrix}$ minimizes $R(\cdot|\Phi,V)$ with respect to

L_{12} and L_{22} at all points (Φ,V), where $\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}' & (I-L_{22})^{-1}L_{22}V_{22} \end{bmatrix}$,

while Φ_{11},Φ_{12} are selected so that Φ is n.n.d.

Because $V_{22}+\Phi_{22}$ is p.d., estimator LY is uniquely determined, and therefore, LY is admissible.

This terminates the proof of sufficiency.

(ii) Now suppose LY is admissible. By Theorem 3.2, there exists a nonzero pair $(\Phi,V) \in [T]$ such that $L(\Phi + V) = \Phi$. If $V \neq 0$, the assertion follows from Lemmas 2.4 and 2.5. If $V = 0$, then $L(\Phi + V) = \Phi$ reduces to $L\Phi = \Phi$, where $\Phi \neq 0$. In this case $\lambda = 1$ is an eigenvalue of L. Without loss of generality we may assume that $L = \begin{pmatrix} I & L_{12} \\ 0 & L_{22} \end{pmatrix}$,

where I is $r \times r$, and $L_{12}Q \neq Q$ for all $Q \in R^{n-r}$ such that $L_{22}Q = Q$. By Corollary 3.3 there exists a point $(V_{12},V_{22},\Phi_{22}) \in \omega$ such that $L_{22}(V_{22} + \Phi_{22}) = \Phi_{22}$, $L_{12}(V_{22} + \Phi_{22}) = -V_{12}$. Because $V_{22} \neq 0$ by Corollary 3.4, it follows that $V_{22} + \Phi_{22}$ is p.d. Hence by Lemma 2.5, all eigenvalues of L_{22} are real and are in $[0,1]$. Moreover, LV is symmetric. (This is shown after Theorem 4.1.)

In the next two sections we shall show that by using the characterization of admissible estimators established in Theorems 3.2 and 3.4, various extensions of Cohen's result to models more general than the ones considered by Cohen himself and by C.R. Rao are possible.

IV. FIRST TWO BASIC THEOREMS

In this section we shall prove two basic theorems -- Theorems 4.1 and 4.2. The first one establishes necessary condition for admissibility of a linear estimator within model with p.s. $\mathcal{P} = R^n \times \mathcal{V}$ without additional assumptions placed on $\mathcal{V}$.

The second theorem may be considered as a generalization of Cohen's theorem to n.n.d. matrices.

First we shall establish two lemmas that are needed to prove the results of this section. As above let Y be an n-variate random vector with parameter space $\mathcal{P} = R^n \times \mathcal{V}$. Consider in addition an (n-r)-variate random vector Z with parameter space $\tilde{\mathcal{P}} = R^{n-r} \times \tilde{\mathcal{V}}$, where $\tilde{\mathcal{V}}$ is a subset of $\mathcal{S}_{(n-r)\times(n-r)}$ defined by $\tilde{\mathcal{V}} = \{V_{22}$: there exist an $r \times r$ matrix V_{11} and an $r \times (n-r)$ matrix V_{12} such that $\begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix} \in \mathcal{V}\}$.

Lemma 4.1.

If LY is admissible within model with p.s. $\mathcal{P} = R^n \times \mathcal{V}$ and if $L = \begin{pmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{pmatrix}$, where L_{22} is $(n-r) \times (n-r)$, then $L_{22}Z$ is admissible within model with p.s. $\tilde{\mathcal{P}} = R^{n-r} \times \tilde{\mathcal{V}}$.

Partition θ into $\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix}$ and V into $\begin{bmatrix}V_{11} & V_{12}\\ V_{12}' & V_{22}\end{bmatrix}$, where θ_1 is $r \times 1$ and V_{11} is $r \times r$. By virtue of Lemma 3.1 we have

$$R(L|\theta,V) = R(L_{22}|\theta_{22},V_{22}) + R_0(L_{11},L_{12},\theta,V),$$

where the last term is nonnegative for all $(\theta,V) \in \mathcal{P}$, $L_{11} \in \mathcal{S}_{r\times r}$ and $L_{12} \in \mathcal{S}_{r\times(n-r)}$. This implies that $L_{22}Z$ is admissible within model p.s. $\tilde{\mathcal{P}}$. In fact, otherwise LY would be inadmissible within model with p.s. $\mathcal{P}$.

Theorem 4.1.

If LY is admissible within model with p.s. $\mathcal{R}^n \times \mathcal{V}$, then

(i) the eigenvalues of L are real and in the closed interval $[0,1]$;

(ii) there exists a nonzero matrix $V \in \mathcal{V}$ such that LV is symmetric.

Proof.

Because LY is admissible, there exists by Theorem 3.2 a nonzero pair $(\Phi,V) \in [\mathcal{T}]$ such that

$$L(\Phi + V) = \Phi. \tag{4.1}$$

Thus $L(\phi + V)$ is symmetric and, therefore, by Lemma 2.1 matrix L has $r = \text{rank}\,(\phi + V)$ independent real eigenvectors such that $\Phi + V = \sum_{i=1}^{r} \tau_i P_i P_i'$, where $\tau_1,\dots,\tau_r$ are some positive numbers. Suppose that $P_1,\dots,P_r$ correspond to eigenvalues $\lambda_1,\lambda_2,\dots,\lambda_r$ of L, respectively. Then

$$(4.2) \qquad \phi = L(V + \phi) = \sum_{i=1}^{r} \tau_i L P_i P_i' = \sum_{i=1}^{r} \tau_i \lambda_i P_i P_i',$$

and, therefore,

$$(4.3) \qquad V = (\phi + V) - \phi = \sum_{i=1}^{r} \tau_i (1 - \lambda_i) P_i P_i'.$$

Because ϕ and V are n.n.d., it follows from (4.2) that $\lambda_i \geq 0$ and from (4.3) that $1 - \lambda_i \geq 0$. Thus $0 \leq \lambda_i \leq 1$, $i = 1,\dots,r$.

To show that the remaining eigenvalues of L are real and in $[0,1]$ when $r < n$, let us choose an orthogonal matrix R such that

$$M = RLR' = \begin{pmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{pmatrix},$$

where M_{11} is $r \times r$ and has eigenvalues $\lambda_1,\dots,\lambda_r$. Moreover, let

$$\tilde{\mathcal{V}} = \{RVR' :\ V \in \mathcal{V}\}.$$

Since LY is admissible within model with p.s. $\mathcal{R}^n \times \mathcal{V}$, estimator MZ is admissible within model with p.s. $\mathcal{R}^n \times \tilde{\mathcal{V}}$. Clearly, Z stands here

for an n-variate random variable with p.s. $R^n \times \tilde{V}$. Now from Lemma 4.1 it follows that $M_{22}X$ is admissible within model with p.s. $R^{n-r} \times \mathcal{U}$, where $\mathcal{U}$ is defined by $\mathcal{U} = \{V_{22}$: there exist an $r \times r$ matrix V_{11} and an $r \times (n-r)$ matrix V_{12} such that $\begin{bmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{bmatrix} \in \mathcal{U}\}$.

It should be obvious that X stands here for an $(n-r)$-variate random variable with p.s. $R^{n-r} \times \mathcal{U}$. In view of the above, we may conclude that M_{22} has at least one eigenvalue in $[0,1]$, which is evidently also an eigenvalue of L. If necessary, we may repeat the reasoning a number of times more to conclude finally that all eigenvalues of L are in $[0,1]$.

We need yet to show that there exists a nonzero matrix V such that LV is symmetric. Clearly, this assertion follows from (4.1) by Lemma 2.4 if V is not the zero matrix. In case V is the zero matrix, (4.1) reduces to $L\Phi = \Phi$. Then as already noted above $\lambda = 1$ is an eigenvalue of L. Suppose that L has an r-fold degeneracy for $\lambda = 1$. Without loss of generality we may then assume that L is of the form

$$L = \begin{bmatrix} I & L_{12} \\ 0 & L_{22} \end{bmatrix}, \quad I \text{ is } r \times r,$$

where $L_{12}Q \neq 0$ if $L_{22}Q = Q$.

Now let $\mathcal{W}$ be defined as in Theorem 3.4. If $\mathcal{W}$ is compact, then by Corollary 3.3 there exists a point

$$(\Phi, V) = \left[\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}' & \Phi_{22} \end{pmatrix}, \begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix}\right] \in [\mathcal{T}]$$

with $(V_{12}, V_{22}, \Phi_{22})$ in $\mathcal{W}$ such that

$$L_{12}(V_{22} + \Phi_{22}) = -V_{12}$$

$$L_{22}(V_{22} + \Phi_{22}) = \Phi_{22}.$$

Recall that $V_{22} = 0$ would contradict the assumption that $L = \begin{pmatrix} I & L_{12} \\ 0 & L_{22} \end{pmatrix}$ has an r-fold degeneracy for $\lambda = 1$. Thus $V_{22} \neq 0$. Because $(V_{12}, V_{22}, \Phi_{22}) \in \mathcal{W}$, there exists an $r \times r$ matrix V_{11} such that $V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix} \in \mathcal{V}$. Now we show that LV is symmetric. Note that

$$LV = \begin{pmatrix} I & L_{12} \\ 0 & L_{22} \end{pmatrix}\begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix} = \begin{pmatrix} V_{11} + L_{12}V_{12}' & V_{12} + L_{12}V_{22} \\ L_{22}V_{12}' & L_{22}V_{22} \end{pmatrix}.$$

To show that LV is symmetric we need to show that

(α) $L_{12}V_{12}'$ is symmetric

(β) $V_{12} + L_{12}V_{22} = V_{12}L_{22}'$

(γ) $L_{22}V_{22}$ is symmetric.

To show (α) note that

$$L_{12}V_{12}' = -L_{12}(V_{22} + \Phi_{22})L_{12}'.$$

To show (β) note that

$$\begin{aligned} V_{12} + L_{12}V_{22} &= -L_{12}\Phi_{22} \\ &= -L_{12}(V_{22} + \Phi_{22})(V_{22} + \Phi_{22})^{+}\Phi_{22} \\ &= V_{12}(V_{22} + \Phi_{22})^{+}\Phi_{22} \\ &= V_{12}(V_{22} + \Phi_{22})^{+}(V_{22} + \Phi_{22})L_{22}' \\ &= V_{12}L_{22}'. \end{aligned}$$

Finally, (γ) follows from Lemma 2.4, because $L_{22}(V_{22} + \Phi_{22}) = \Phi_{22}$.

Now we shall consider the case when W is not compact. Then there exist some sequences $\{\Phi_{22}^{(n)}\}$ and

$$V^{(n)} = \begin{pmatrix} V_{11}^{(n)} & V_{12}^{(n)} \\ V_{12}'^{(n)} & V_{22}^{(n)} \end{pmatrix}$$

such that

$$X_n = (V_{12}^{(n)}, V_{22}^{(n)}, \Phi_{22}^{(n)}) \in \omega$$

and

$$X_n \to X_0 = (V_{12}^{(0)}, V_{22}^{(0)}, \Phi_{22}^{(0)}) \notin \omega.$$

Because no subsequence of $\{V_{11}^{(n)}\}$ may converge to a n.n.d. matrix, the elements of $\{V_{11}^{(n)}\}$ are not bounded. In this case there exists a subsequence $\{n_i\}$ such that

$$\frac{1}{\operatorname{tr} V^{(n_i)}} V^{(n_i)} \to \begin{pmatrix} W_{11} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{as } i \to \infty,$$

where $W_{11} \neq 0$. Because

$$\begin{pmatrix} I & L_{12} \\ 0 & L_{22} \end{pmatrix} \begin{pmatrix} W_{11} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} W_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

is symmetric, the proof is terminated.

Now let $\mathcal{V}$ be a closed convex cone of $n \times n$ n.n.d. matrices defined by

$$\mathcal{V} = \{\begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix} : W \in \mathcal{U}\},$$

where $\mathcal{U}$ may stand for an arbitrary closed convex cone of $r \times r$ n.n.d. matrices, where $1 \leqslant r < n$. As above, consider an n-variate random vector Y with p.s. $\mathcal{R}^n \times \mathcal{V}$ and in addition, an r-variate random vector Z with p.s. $\mathcal{R}^r \times \mathcal{U}$.

We shall characterize admissible estimators within model with p.s. $\mathcal{R}^n \times \mathcal{V}$, where $\mathcal{V}$ has the structure described above. First we prove the following lemma.

Let $L = \begin{pmatrix} L_{11} & L_{12} \\ 0 & I \end{pmatrix}$ and let $L_0 = \begin{pmatrix} L_{11} & (I - L_{11})H \\ 0 & I \end{pmatrix}$, where H may be an arbitrary $r \times (n-r)$ matrix, such that $(I - L_{11})'L_{12} = (I - L_{11})'(I - L_{11})H$. Such a matrix H always exists, because

$$R((I - L_{11})') \subset R((I - L_{11})'(I - L_{11})).$$

Lemma 4.2.

Within model with p.s. $R^n \times V$ estimator L_0Y is as good as LY and better than LY when $R(L_{12}) \not\subset R(I - L_{11})$.

Proof.

Note that

$$R(L|\theta,V) = \operatorname{tr} L_{11}'L_{11}W + \theta'\begin{pmatrix} I - L_{11}' \\ -L_{12}' \end{pmatrix}(I - L_{11}, -L_{12})\theta$$

and that

$$R(L_0|\theta,V) = \operatorname{tr} L_{11}'L_{11}W + \theta'\begin{pmatrix} I - L_{11}' \\ -H'(I - L_{11})' \end{pmatrix}(I - L_{11}, -(I - L_{11})H)\theta .$$

Denote by $(I - L_{11})^{\perp}$ a matrix such that

$$R((I - L_{11})^{\perp}) = N(I - L_{11}).$$

Then there exists a matrix S such that

$$L_{12} = (I - L_{11})H + (I - L_{11})^{\perp}S.$$

Consequently,

$$(4.4) \qquad L_{12}'L_{12} = H'(I - L_{11}')(I - L_{11})H + S'[(I - L_{11})^{\perp}]'(I - L_{11})^{\perp}S.$$

Because

$$\begin{pmatrix} I - L_{11}' \\ -L_{12}' \end{pmatrix} (I - L_{11}, -L_{12}) = \begin{pmatrix} (I - L_{11})'(I - L_{11}) & -(I - L_{11})'L_{12} \\ -L_{12}'(I - L_{11}) & L_{12}'L_{12} \end{pmatrix}$$

and because

$$\begin{pmatrix} I - L_{11}' \\ -H'(I - L_{11})' \end{pmatrix} (I - L_{11}, -(I - L_{11})H = \begin{pmatrix} (I - L_{11})'(I - L_{11}) & -(I - L_{11})'(I - L_{11})H \\ -H'(I - L_{11})'(I - L_{11}) & H'(I - L_{11})'(I - L_{11})H \end{pmatrix}$$

we note that

$$R(L|\theta,V) - R(L_0|\theta,V) = \theta' \begin{pmatrix} 0 & 0 \\ 0 & L_{12}'L_{12} - H'(I - L_{11})'(I - L_{11})H \end{pmatrix} \theta .$$

Thus in order that L_0Y be as good as LY it is sufficient that

$$L_{12}'L_{12} \geq H'(I - L_{11})'(I - L_{11})H.$$

However this follows from (4.4), because

$$[(I - L_{11})^{\perp}S]'(I - L_{11})^{\perp} S \geq 0.$$

This terminates the proof.

Now we are able to prove the following result.

Theorem 4.2.

Estimator LY, where

$$L = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}, \quad L_{11} \text{ is } r \times r,$$

is admissible within model p.s. $\mathcal{R}^n \times \mathcal{V}$ if and only if

(i) $L_{21} = 0$, $L_{22} = I$, $\mathcal{R}(L_{12}) \subset \mathcal{R}(I - L_{11})$,

and

(ii) $L_{11}Z$ is admissible within model with p.s. $\mathcal{R}^r \times \mathcal{U}$.

First suppose that LY is admissible. Because $\begin{pmatrix} L_{11} & L_{12} \\ 0 & I \end{pmatrix}Y$ is better than $\begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}Y$ when $L_{21} \neq 0$ and/or $L_{22} \neq I$, we must have $L_{21} = 0$ and $L_{22} = I$. Also there exists a matrix H such that $L_{12} = (I - L_{11})H$. Otherwise $\begin{pmatrix} L_{11} & (I - L_{11})H \\ 0 & I \end{pmatrix}Y$, where H is such that $(I - L_{11})'L_{12} = (I - L_{11})'(I - L_{11})H$, would be better than $\begin{pmatrix} L_{11} & L_{12} \\ 0 & I \end{pmatrix}Y$.

Hence $R(L_{12}) \subset R(I - L_{11})$. Now suppose that $L_{11}Z$ is not admissible within model with p.s. $R^r \times \mathcal{U}$. Let $M_{11}Z$ be better. Then for all $W \in \mathcal{U}$

$$\text{(4.5)} \qquad \operatorname{tr} M_{11}'M_{11}W \leq \operatorname{tr} L_{11}'L_{11}W$$

and

$$\text{(4.6)} \qquad (I - M_{11})'(I - M_{11}) \leq (I - L_{11})'(I - L_{11})$$

and

$$\text{(4.7)} \qquad R(M_{11}|\theta_1,W) \leq R(L_{11}|\theta_1,W)$$

for all $(\theta,W) \in R^r \times \mathcal{U}$ and strict inequality for at least one point in $R^n \times \mathcal{U}$.

Now from (4.6) it follows that

$$\begin{pmatrix} I \\ -H' \end{pmatrix}(I - M_{11})'(I - M_{11})(I,-H) \leq \begin{pmatrix} I \\ -H' \end{pmatrix}(I - L_{11})'(I - L_{11})(I,-H).$$

Hence for

$$M = \begin{pmatrix} M_{11} & (I - M_{11})H \\ 0 & I \end{pmatrix}$$

we have

$$R(M|\theta,V) \leq R(L|\theta,V) \quad \text{for all} \quad (\theta,V) \in R^n \times \mathcal{V}.$$

Because LY is admissible, it follows that

$$R(M|\theta,V) = R(L|\theta,V).$$

This implies that for all $W \in \mathcal{U}$

$$\mathrm{tr}\, M_{11}'M_{11}W = \mathrm{tr}\, L_{11}'L_{11}W$$

and that

$$(I - M_{11})'(I - M_{11}) = (I - L_{11})'(I - L_{11})$$

which contradict assumption (4.7) that $M_{11}Z$ is better than $L_{11}Z$. Thus $L_{11}Z$ is admissible.

Now suppose that L meets conditions (i) and (ii) of the theorem. And suppose that LY is not admissible. Let MY, where

$$M = \begin{pmatrix} M_{11} & (I - M_{11})S \\ 0 & I \end{pmatrix},$$

be better than LY.

Now

$$R(M|\theta,V) = \mathrm{tr}\, M_{11}'M_{11}W + \theta'\begin{pmatrix} I - M_{11}' \\ -S'(I - M_{11})' \end{pmatrix}(I - M_{11}, (I - M_{11})S)\theta.$$

If $R(M|\theta,V) \leq R(L|\theta,V)$, then

$$\mathrm{tr}\, M_{11}'M_{11}W + \theta_1'(I - M_{11})'(I - M_{11})\theta_1 \leq \mathrm{tr}\, L_{11}'L_{11}W + \theta_1'(I - L_{11})'(I - L_{11})\theta_1.$$

Because $L_{11}Z$ is admissible by assumption, therefore we may replace $\leq$ by

=. But in this case

$$0 \leqslant R(M|\theta,V) - R(L|\theta,V)$$
$$= \theta' \begin{pmatrix} C_{11} & C_{12} \\ C'_{12} & C_{22} \end{pmatrix} \theta,$$

where

$$C_{11} = 0$$

$$C_{12} = -(I - M)'(I - M_{11})S + (I - L_{11})'(I - L_{11})H$$

$$C'_{12} = -S(I - M_{11})'(I - M_{11}) + H'(I - L_{11})'(I - L_{11})$$

and

$$C_{22} = S'(I - M_{11})'(I - M_{11})S - H'(I - L_{11})'(I - L_{11})H$$

so that

$$S'(I - M_{11})'(I - M_{11})S = H'(I - L_{11})'(I - L_{11})H$$

by using

$$(I - M_{11})'(I - M_{11}) = (I - L_{11})'(I - L_{11}).$$

Consequently, $R(M|\theta,V) = R(L|\theta,V)$ contrary to the assumed. This contradiction proves that LY is admissible.

In case $V = [V_0]$, where $V_0 = \begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix}$, while W is a p.d. matrix of order $r \times r$, the following corollary follows immediately from Theorems 4.2 and 3.5. Let $L = \begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{pmatrix}$, where L_{11} is $r \times r$.

Corollary 4.1.

Estimator LY is admissible with model with p.s. $R^n \times V$ if and only if

(i) $L_{21} = 0$, $L_{22} = I$, $R(L_{12}) \subset R(I - L_{11})$

and

(ii) $L_{11}W$ is symmetric and the eigenvalues of L_{11} are in the closed interval $[0,1]$.

V. THREE MORE THEOREMS ON ADMISSIBILITY

In this section we show first that all linear estimators admissible within model $R^n \times \mathcal{V}_1$ are also admissible within model $R^n \times \mathcal{V}_2$ if $\mathcal{V}_1 \subset \mathcal{V}_2$. Next we show that the conditions appearing in Theorem 4.1 are sufficient for LY to be admissible when $\mathcal{V}_1$ consists of p.d. matrices only. Finally, we establish the main result of the paper that if $\mathcal{V}$ is the set of all n.n.d. matrices, then LY is admissible if and only if all the eigenvalues of L are in $[0,1]$.

The essential tool to prove the above theorems is the following easy-to-prove result.

Lemma 5.1.

If LY is admissible within model with p.s. $\mathcal{P} = R^n \times \mathcal{V}$ and if MY is another estimator with the same risk function, i.e., $R(M|\theta,V) = L(L|\theta,V)$ for all $(\theta,V) \in \mathcal{P}$, then $M = L$.

Proof.

By Lemma 2.6 the risk function $R(L|\theta,V) = \mathrm{tr}\, L'LV + \mathrm{tr}\,(I-L)'(I-L)\theta\theta'$ is a convex function of L for all $(\theta,V) \in \mathcal{P}$ such that V is p.d. Suppose that there exists a p.d. matrix, say V_0, in $\mathcal{V}$. If $T = \frac{1}{2}(L+M)$, then in view of the above $R(T|\theta,V) \leq R(L|\theta,V)$ for all $(\theta,V) \in \mathcal{P}$ with a strict inequality for $V = V_0$ unless $L = M$, and this contradicts the admissibility of L. In case all matrices in $\mathcal{V}$ are singular we may assume

without loss of generality that $\mathcal{U} = \left\{ \begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix} : W \in \mathcal{U} \right\}$, where not all matrices in $\mathcal{U}$ are singular. By Theorem 4.2, matrices L and M must then be necessarily of the following form: $L = \begin{pmatrix} L_{11} & L_{12} \\ 0 & I \end{pmatrix}$ and $M = \begin{pmatrix} M_{11} & M_{12} \\ 0 & I \end{pmatrix}$, where $\mathcal{R}(L_{12}) \subset \mathcal{R}(I - L_{11})$ and $\mathcal{R}(M_{12}) \subset \mathcal{R}(I - M_{11})$. Moreover, by Theorem 4.2, both $L_{11}Z$ and $M_{11}Z$ are admissible within model with p.s. $R^r \times \mathcal{U}$, where $1 \leq r < n$. Hence $M_{11} = L_{11}$, since $L_{11}Z$ and $M_{11}Z$ have the same risk function. From this and the assumption that LY and MY have the same risk function, we obtain that $(I - L_{11})'(M_{12} - L_{12}) = 0$ which implies that $M_{12} = L_{12}$. Thus $L = M$ and the proof is terminated.

A typical situation where Lemma 5.1 may be used is the following one. Consider two models with p.s. $P_1 = R^n \times V_1$ and $P_2 = R^n \times V_2$, say. Assume that $V_1 \subset V_2$ and that LY_1 is admissible with model with p.s. P_1. The question is whether LY_2 is admissible within model with p.s. P_2. To answer this question suppose that MY_2 is as good as LY_2 within model P_2 and, consequently, also within model P_1. Hence $R(M|\theta,V) = R(L|\theta,V)$ for all $(\theta,V) \in P_1$ because LY_1 is admissible within P_1. Now $M = L$ follows from Lemma 5.1 so that the following result holds.

Theorem 5.1.

Each admissible linear estimator within model with p.s. $R^n \times V_1$ is also admissible within model with p.s. $R^n \times V_2$ if $V_1 \subset V_2$.

Using Lemma 5.1 and Theorem 3.5 we also arrive straightforward at Theorem 5.2 below which asserts that the conditions in Theorem 4.1 are sufficient in case V consists of p.d. matrices only.

Theorem 5.2.

Let V be a closed convex cone of p.d. matrices. Then LY is admissible within model with parameter space $R^n \times V$ if and only if the eigenvalues of L are in [0,1] and there exists a nonzero matrix V such that LV is symmetric.

Proof.

We need only show that if the eigenvalues of L are in [0,1] and if LV_0, where $V_0 \in V$, $V_0 \neq 0$, is symmetric, then LY is admissible within model with p.s. $R^n \times V$.

Suppose that LY is not admissible within the model considered. Then there exists a better estimator than LY. Suppose that MY is better than LY. Then

$$R(M|\theta,V) \leq R(L|\theta,V) \quad \text{for all} \quad (\theta,V) \in R^n \times V.$$

In particular,

$$R(M|\theta,V_0) \leq R(L|\theta,V_0) \quad \text{for all} \quad \theta \in R^n.$$

Since LV_0 is symmetric and since the eigenvalues of L are in [0,1], it

follows from Theorem 3.5, that LY is admissible with model with p.s. $R^n \times [V_0]$. Thus for all $\theta \in R^n$ holds

$$R(M|\theta, V_0) = R(L|\theta, V_0).$$

Now we may apply Lemma 5.1 and conclude that $M = L$. This terminates the proof.

The following interesting corollary follows straightforwardly from Theorem 5.2 and Theorem 3.5. Let $[V]$ denote the closed convex cone generated by n.n.d. matrix V.

Corollary 5.1.

Under the assumptions of Theorem 5.2 estimator LY is admissible if and only if there exists a nonzero matrix $V \in \mathcal{V}$ such that LY is admissible within model with p.s. $R^n \times [\mathcal{V}]$.

The following example illustrates that the assertion of Corollary 5.1 is false when $\mathcal{V}$ contains a singular matrix.

Example 5.1.

Let $\mathcal{V}$ be the closed convex cone generated by matrices $V_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\mathcal{V}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then $V = \{\tau \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} : \alpha \in [0,1],\ \tau \geq 0\}$.

First we show that LY, where $L = \begin{pmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix}$, is admissible within model with p.s. $R^2 \times \mathcal{V}$.

Observe that L meets conditions (3.2) when $V = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\Phi = \begin{pmatrix} 16 & 4 \\ 4 & 1 \end{pmatrix}$. On the other hand, if MY is as good as LY, then M must be of the form $M = \begin{pmatrix} \frac{1}{2} & m_{12} \\ 0 & m_{22} \end{pmatrix}$ by Corollary 3.1. Moreover, by Theorem 3.3, matrix M meets conditions (3.2) for $V = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\Phi = \begin{pmatrix} 16 & 4 \\ 4 & 1 \end{pmatrix}$ too, i.e.,

$$m_{12}(1 + 1) = 2$$

$$m_{22}(1 + 1) = 1.$$

Hence $m_{12} = 1$ and $m_{22} = 1$, and consequently, $M = L$. This shows that LY is admissible within the specified model.

Now suppose that there exists a matrix $V \in \mathcal{V}$ such that LY is admissible within model with p.s. $R^2 \times V$. Then, by Theorem 4.1, matrix LV must be symmetric. Now LV is symmetric if and only if $V = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, because $LV = \begin{pmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \alpha \\ 0 & \frac{1}{2}\alpha \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ \alpha & \frac{1}{2}\alpha \end{pmatrix}$.

Now observe that LY, where $L = \begin{pmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix}$, is not admissible within model

with p.s. $\mathcal{R}^2 \times [\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}]$. As already established LY, where

$L = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix}$, is admissible within model with p.s. $\mathcal{R}^2 \times [\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}]$

if and only if $l_{21} = 0$, $l_{22} = 1$ and $l_{12} = (1 - \frac{1}{2})h$, $h \in \mathcal{R}$, by Corollary 4.1.

Matrix L does not meet these conditions.

Theorem 5.3.

Let $\mathcal{V}_n$ be the set of all n.n.d. matrices of order $n \times n$. Then LY is admissible within model with p.s. $\mathcal{R}^n \times \mathcal{V}_n$ if and only if all eigenvalues of L are real and are in $[0,1]$.

To prove this theorem we need the following result.

Lemma 5.2.

Let L be of the form $L = \begin{pmatrix} \lambda & L_{12} \\ 0 & L_{22} \end{pmatrix}$, where $\lambda \in \mathcal{R}$. Then LY is admissible within model with p.s. $\mathcal{R}^n \times \mathcal{V}_n$ if and only if $L_{22}Z$ is admissible within model with p.s. $\mathcal{R}^{n-1} \times \mathcal{V}_{n-1}$ and $\lambda \in [0,1]$.

Here Z stands for an (n-1)-variate random variable with p.s. $\mathcal{R}^{n-1} \times \mathcal{V}_{n-1}$.

Proof.

If LY is admissible within model with p.s. $\mathcal{R}^n \times \mathcal{V}_n$, then $\lambda \in [0,1]$ by Theorem 4.1. Moreover, $L_{22}Z$ is admissible within model with p.s. $\mathcal{R}^{n-1} \times \mathcal{V}_{n-1}$ by Lemma 4.1.

To show the converse, suppose that the assertion is false. In other words, assume that there exists an estimator MY such that

$$R(M|\theta,V) \leqslant R(L|\theta,V)$$

for all $(\theta,V) \in \mathcal{R}^n \times \mathcal{V}_n$ and strict inequality for at least one point in $\mathcal{R}^n \times \mathcal{V}_n$. Because $LP = \lambda P$ for $P = (1,0,\ldots,0)'$ we note that $MP = \lambda P$ by Corollary 3.1. Hence

$$M = \begin{pmatrix} \lambda & M_{12} \\ 0 & M_{22} \end{pmatrix} .$$

Because

$$\begin{aligned} R(L|\theta,V) &= \operatorname{tr} \begin{pmatrix} \lambda & 0 \\ L_{12}' & L_{22}' \end{pmatrix} \begin{pmatrix} \lambda & L_{12} \\ 0 & L_{22} \end{pmatrix} \begin{pmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{pmatrix} \\ &\quad + (\theta_1' \ \theta_2') \begin{pmatrix} 1-\lambda & 0 \\ -L_{12}' & I - L_{22}' \end{pmatrix} \begin{pmatrix} 1-\lambda & -L_{12}' \\ 0 & I - L_{22} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \\ &= \lambda^2 V_{11} + 2\lambda L_{12} V_{12}' + \operatorname{tr} (L_{12}'L_{12} + L_{22}'L_{22})V_{22} \\ &\quad + (\theta_1' \ \theta_2') \begin{pmatrix} (1-\lambda)^2 & -(1-\lambda)L_{12} \\ -(1-\lambda)L_{12}' & L_{12}'L_{12} + (I - L_{22}')(I - L_{22}) \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \end{aligned}$$

we have for all $(\theta, V) \in R^n \times V_n$

$$(5.1) \quad R(L|\theta,V) - R(M|\theta,V) = R(L_{22}|\theta_2,V_{22}) - R(M_{22}|\theta_2,V_{22})$$
$$+ \operatorname{tr}(L_{12}'L_{12} - M_{12}'M_{12})(V_{22} + \theta_2\theta_2')$$
$$+ 2\lambda(L_{12} - M_{12})V_{12}' - 2(1-\lambda)(L_{12} - M_{12})\theta_2\theta_1$$
$$\geq 0.$$

Note that θ_1 appears in the last term only.

Since θ_2 and θ_1 may be arbitrary, inequality (5.1) implies that $L_{12} = M_{12}$ in case $\lambda \neq 1$. On the other hand, if $\lambda = 1$, then (5.1) reduces to

$$(5.2) \qquad R(L_{22}|\theta_2,V_{22}) - R(M_{22}|\theta_2,V_{22}) = \operatorname{tr}(L_{12}'L_{12} - M_{12}'M_{12})(V_{22} + \theta_2\theta_2')$$
$$+ 2(L_{12} - M_{12})V_{12}'$$
$$\geq 0.$$

Since V_{11} does not appear in (5.2) and since V_{12} may be arbitrary, provided $R(V_{12}') \subset R(V_{22})$, we conclude from (5.2) that M_{12} must be equal to L_{12}, when (5.1) holds for all $(\theta,V) \in R^n \times V_n$. Therefore, in either event (5.1) reduces to $R(M_{22}|V_{22},\theta_2) \leq R(L_{22}|V_{22},\theta_2)$ from which we conclude that $M_{22} = L_{22}$ by Lemma 5.1. Thus, indeed $M = L$, and this contradicts the assumption that MY is better than LY. Now we are in a position to establish Theorem 5.3.

Proof of Theorem 5.3.

We need only show that if the n eigenvalues $\lambda_1,\ldots,\lambda_n$ of L are in $[0,1]$, then LY is admissible within the specified model. The converse is guaranteed by Theorem 4.1.

Let P_1 be an eigenvector of L with eigenvalue λ_1. Assuming that $P_1 = (1,0,\ldots,0)'$, matrix L must then be of the following form

$$L = \begin{pmatrix} \lambda_1 & L_{12} \\ 0 & L_{22} \end{pmatrix}.$$

To prove that LY is admissible within model with p.s. $R^n \times V_n$, it suffices to show by Lemma 5.2 that $L_{22}Z$ is admissible within model with a reduced p.s. $R^{n-1} \times V_{n-1}$. Because $\lambda_2,\ldots,\lambda_n$ are eigenvalues of L_{22}, verification of admissibility of $L_{22}Z$ may be reduced similarly as above to verification of admissibility of a linear estimator in a model with $R^{n-2} \times V_{n-2}$. Clearly, continuing this reasoning we are finally led to a 1×1 matrix (λ_n), where $\lambda_n \in [0,1]$. And $\lambda_n Z$, where Z is a random variable with expectation $\Theta = EZ \in R$ and covariance $\sigma^2 I = \text{cov } Z \in R^+$, is admissible. Thus, the original estimator LY is admissible as asserted. This completes the proof.

REFERENCES

[1] Cohen, A. "All admissible linear estimates of the mean vector," Ann. Math. Statist. 37(1966), 458-463.

[2] La Motte, L.R. "Some results on biased estimation applied to variance components estimation." Presented at the VI International Conference on Mathematical Statistics in Wisla, Poland, 1978.

[3] Olsen A., Seely J. and D. Birkes. "Invariant q quadratic unbiased estimation for two variance components," Ann. Stat. 4(1976), 878-890.

[4] Rao, C.R. "Estimation of parameters in a linear model," Ann. Stat. 4(1976), 1023-1037.

[5] Shinozaki, N. "A study of generalized inverse of matrix and estimation with quadratic loss." Ph.D. thesis submitted to the Keio University, Japan, 1975.

DATE
ILMED
-8